

# ディープラーニング時代における 医用画像のコンピュータ支援診断\*

Computer-Aided Diagnosis of Medical Images in the Deep Learning Era

藤田広志\*\*  
Hirosi FUJITA

## Key words

artificial intelligence (AI), deep learning, machine learning, computer-aided diagnosis (CAD), medical application, medical imaging

## 1. はじめに

人工知能 (AI) におけるディープラーニング (deep learning: 深層学習) の出現により, 画像認識精度が人間を超えるレベルに達するものも出現している. 医療分野への AI・ディープラーニングの導入も急激に進み, すでに商用化事例も国内外で増え始めている. 本稿では, 医用画像診断領域における AI (ディープラーニング) 導入の現状や課題について, その歴史も含めて解説する. なお, 本稿は文献 1) を元に加筆・修正したものである. また, 紙幅の制限で概要を解説するに留めるが, より詳細な内容に興味ある読者は, 文献 2)-4)などを参照されたい.

## 2. コンピュータ支援診断: CAD

医用画像診断において, あってはならないが「誤診」はどうしても避けられない課題である. 特に集団検診のように大量に画像が発生するところでは, 誤診が起りやすく, そのため医師からはコンピュータによる診断支援への要望が昔から強く, 昨今のように「画像の洪水」が起きている現代では, ますますその期待が膨らんでいる. また, 本邦は, 対人口比の CT・MRI 保有台数で世界 1 位を誇っている画像診断大国であるにもかかわらず<sup>5)</sup>, 読影 (画像診断) をする放射線科の医師はまったく不足で, したがって医師 1 人当たりの画像の読影件数も世界 1 位という苦しい実情がある.

コンピュータによる画像診断への取り組みには, 「自動診断」 (automated diagnosis, autonomous diagnosis) あるいは「支援診断」 (computer-aided diagnosis, CAD) を目指すとして, すでに 1960 年頃からの長い研究開発の歴史がある. その詳細は文献 2) に譲る. AI の発展の歴史と CAD の進化について図 1 に示す.

自動診断はハードルが相当に高いため, まず実用化に成

功したのは, 乳がん検出を対象としたコンピュータ支援検出 (computer-aided detection, CADe) と呼ばれるシステムである. これは, シカゴ大学で開発されたシステムをベースにして, 米国のベンチャー企業 R2 Technologies 社 (現 Hologic) が, 乳がんの X 線検査に使用される画像 (マンモグラフィと呼ばれる) を対象に開発したもので, 1998 年に米国の食品医薬品局 (food and drug administration, FDA) の審査・承認を得て, 商用化に至っている. 当時はまだアナログ (連続情報) と称せられるフィルム画像が使われていた時代であり, レーザデジタルタイザによってフィルムをスキャンし, コンピュータで取り扱えるデジタルデータ (離散情報) に処理した上で, 画像入力を行う必要があった. この CAD の商用化がスタートした 1998 年は, しばしば「CAD 元年」とも呼ばれる.

その後, このマンモグラフィ CAD に続いて, 肺がん検出を目的とした胸部単純 X 線写真や胸部 CT 画像の CAD が米国で商用化され, さらに大腸がん検査で用いられる CT コロノグラフィ検査で大腸ポリープを自動検出するための CAD も後に続いた.

しかしながら, 実用化 (= 商用化) され多大な普及に結び付いた成功例は, マンモグラフィ CAD に限定される. その最大の成功要因は, マンモグラフィ CAD の利用によ

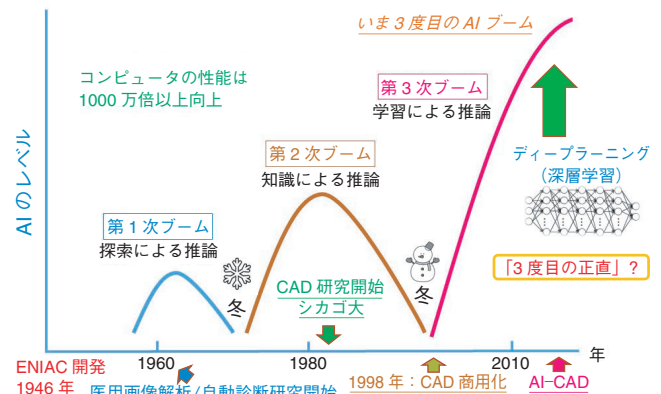


図 1 AI のレベルとその進化, および CAD の開発史

\*原稿受付 令和 3 年 5 月 13 日

\*\*岐阜大学工学部 (岐阜市柳戸 1-1)

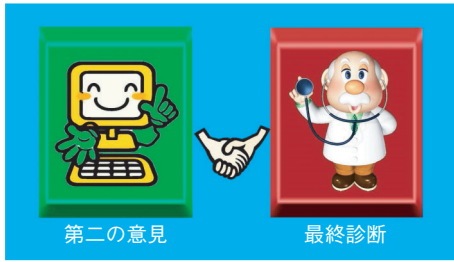


図2 コンピュータ支援診断 (CAD) の概念図

り、2000年初頭から医療保険の補助金が受けられるようになったことである。2016年には検診マンモグラフィの読影の92%でCADが使われるようになっている<sup>6)</sup>。

このようなCADの利用法には、厳密な定義があった。それはコンピュータの結果を「第二の意見」として医師が利用するもので、「セカンドリーダー型 (second reader) CAD」といわれる。すなわち、

- ① 医師はまずCADの結果なしで画像を単独で読影し、その後、
- ② CADの結果(「第二の意見」)を参照して総合的な最終診断を下す、

というものであった(図2)。これには、最初からCADの指摘箇所のみを参照してほかの箇所を見(診)なくなることを防ぎ(CADも見落としをするため)、あるいは拾いすぎ(偽陽性)候補を医師が精査する機会を与えるという効果がある。しかし、一方で読影時間が増加してしまう傾向が指摘された。

CAD商用化歴の約20年の間に、いろいろな問題点も明らかになってきた。

- 1) 開発コストが大きい、
- 2) 検出性能(偽陽性候補も問題)がまだ不十分で、マンモグラフィ検診では要精検率が増加する、
- 3) 商用化に至ったのはすべて検出支援型のCADeであり、検出部位ががんかどうかを識別(鑑別)するような診断支援型のCADx(狭義のcomputer-aided diagnosisの略で、Dxはdiagnosisの略)は、FDAの承認例は皆無、
- 4) 実臨床で大規模検証したところ有効性がなかった、とするいくつかの論文の出現(Fenton論文など)<sup>7)</sup>、
- 5) 特定の病変にしか対応できていない、
- 6) 現場で使い方が面倒!(ワークフローが悪い)、などの理由が挙げられる。

### 3. ディープラーニングによりAI-CADの時代に

2012年の一般画像認識のコンペティション(ILSVRC)で、カナダのトロント大学のHintonらは、ディープラーニングによって1年前の優勝記録の誤認識率を25.7%から15.3%へと4割も削減させ圧勝した。このHintonが、2016年秋にトロントで開催された機械学習に関する国際会議で、「5年以内にディープラーニングは専門医のレベ

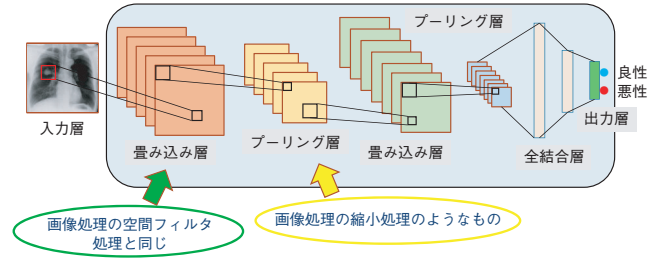


図3 代表的なディープラーニングである畳み込みニューラルネットワーク (CNN) による良悪性鑑別への応用

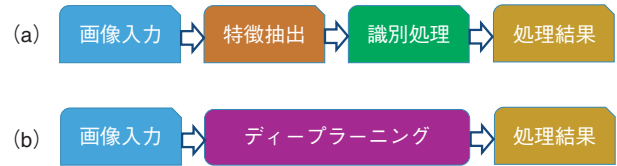


図4 従来型CAD (a) とディープラーニング型AI-CAD (b) の開発過程の相違

ルに達するだろう」という主旨の衝撃的なスピーチをし<sup>8)</sup>、大きな話題となった。

この予言を実証するかのように、まずは放射線画像ではないが、2017年1月に米国のスタンフォード大学のEstevaらが、皮膚がんの診断にディープラーニングを活用した結果をNature誌に報告している<sup>9)</sup>。彼らは約13万枚の皮膚疾患の画像を収集し、「メラノーマ(悪性黒色腫)」や「良性腫瘍」をAIで学習させた結果、皮膚科医と同等の精度で皮膚がんを診断できたという。さらに、その翌年、ドイツのハイデルベルク大学のHaenssleらの「人間 vs. 機械」との主題の論文では、10万枚以上の画像で開発されたディープラーニングと皮膚科医がメラノーマを鑑別したところ、その精度は皮膚科医 vs. AIが87% vs. 95%となり、AIに軍配があがっている<sup>10)</sup>。

いまや伝統的(traditional)とも呼ばれる前項で説明したような「従来型CAD」では、画像の中の認識対象の特徴量を、設計者(人間)が苦勞して考案・作成したのに対して、ディープラーニングの利点は自ら特徴量を作り出す(学習する)ことができる点にある(図3)。すなわち、従来型CADは図4(a)のようであり、ディープラーニングを使うAI-CADは図4(b)のようになり、開発工程が一変し、また、著しい性能向上が見られるケースが多くなった。

どう定義するか次第であるが、旧来型AIにおけるルールベース法に基づくものや、ディープラーニング以外の従来の機械学習を用いるもの(分類・識別処理に人工ニューラルネットワークやサポートベクトルマシンを使う部類)までを「従来型CAD」と呼び、ディープラーニング型のCADを「AI-CAD」と呼んで区別する。

AIのレベルは、図5のように5段階で区別される。ここで、レベル1のAIは、AI搭載をうたった家電製品で



図5 AIの5つのレベル分類

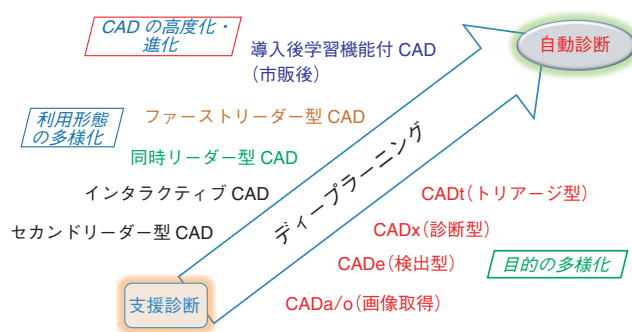


図7 ディープラーニングによる AI-CAD の進化・多様化

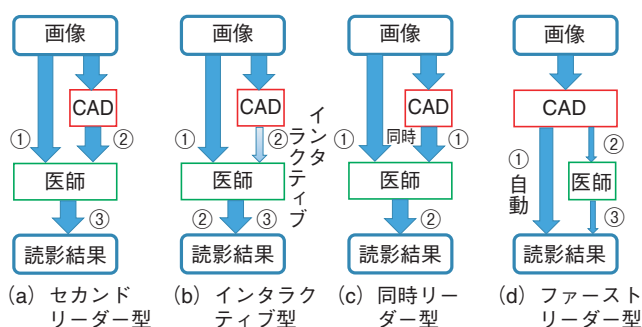


図6 AI-CADの4つの利用形態

(単純な制御プログラム)、レベル2は質問応答システム、お掃除ロボットや一般的な将棋ソフトウェアのレベルである(「知識」を使ったAIで、推論・探索が可能)。これらは、古典的なAIとも呼ばれる。レベル3になると、検索エンジンやビッグデータ分析に活用されるもので、機械学習が行われるようになり、人間から特徴量を教えられて学習する。レベル4では、さらに高度な分析が可能になり、ディープラーニングが取り入れられ、人間が特徴量を教えなくても自力で獲得(学習)する。車の自動運転や、昨今の囲碁AIなどはこれに該当する。しかし、これらのレベルのAIは、特定の目的に限定されたもので、「特化型AI」に分類される。そして、レベル5になると、人間のよう(あるいはそれ以上に)何でもできるAIで、「汎用AI」(artificial general intelligence, AGI)と呼ばれるが、これはまだ実現されていない。レベル5のAIは「強いAI」とも呼ばれ、それ以外は「弱いAI」とも呼ばれることがある。ディープラーニングの出現により、AI(ここではCAD)はレベル4に上がっている。

#### 4. CAD 利用形態の進化・多様化

1998年にFDA承認のマンモグラフィCADが出現して以降、そのすべての利用形態は上述の「セカンドリーダー型」のみであった。ところが、2016年以降、これに変化が起きている(図6)。また、利用形態の多様化のみなら

ず、その高度化や利用目的の多様化も起きている(図7)。詳細は文献1)~4)に譲るが、両図中で、いまだに商用化に至っていないのは、「ファーストリーダー型CAD」と「導入後学習機能付きCAD」である。前者は、最初にCADが単独で解析処理を行い、医師がチェックすべき画像と明らかに正常でその必要がないものを選定し、その画像の解析結果も提示するタイプのCADである。これは半自動診断に相当し、ほとんどが正常症例である検診での利用が強く望まれる。後者は、臨床現場に導入後(市販後)にも、新しいデータの追加・再学習により、プログラムがアップデートされ、「どんどん賢くなるCAD」であるが、逆に性能劣化の危険性もあり、まだ実用化には課題が多い。

CADの目的も従来の検出型のCADeから多様化が起きており、FDA承認を取得し商用化に成功を収めているものが少なからず出現している。がんの良悪性鑑別を行う診断型のCADxやトリアージ(優先順位付)を目的としたトリアージ型のCADtである。

話題になっているのは、眼底写真から糖尿病網膜症を自動診断する医療AI機器が、2018年にFDAで認可を得ていることである。ただ、2020年に米国の2つの医学放射線学会が、このような自律型システムや導入後学習機能付CADの承認は、安全性の面から時期尚早として、連邦政府に承認のブレーキを掛けるように要請書を提出している。

なお、AI-CADの本格的な普及にはいわゆる保険の償還が重要であるが、米国では最近2つの医療AI機器(脳卒中を警告するCADtと上記の眼底診断に関するもの)が成功しており、今後、さらにそのようなケースが増えることが期待される。

#### 5. 医用画像データについて

すでに多くの研究事例が示しているように、ディープラーニングのCADへの影響力は計り知れないが、ディープラーニング型のAI-CADはデータ駆動型(data driven)であるため、勝負は画像データの収集パワーに依存する要素が大である(図8)。「量」が重要ではあるが、「Garbage in, garbage out」といわれるように、その「質」



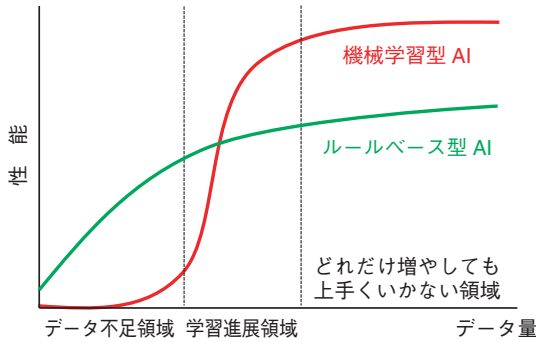


図 8 性能とデータ量

研究における障壁

1. 学習データ不足  
⇒ 各種領域での公共の画像データベースのさらなる開発  
⇒ 少数の教師データ、あるいは教師なし学習法の開発
2. 実験室研究・検討から実臨床研究・検証へ

AIの実装における課題

3. AIの効果的な利用法(支援診断, 半自動診断, 自動診断)
4. 既存のシステムへの組み込みとシームレス化
5. 薬事承認と保険償還

倫理的側面

6. AI診断に対する責任の所在(現在は医師に)  
⇒ 医師? 病院? ベンダー?
7. ディープラーニングのブラックボックス性の解決  
⇒ 説明可能な AI(explainable AI) 研究へ
8. 医療関係者から患者までの AI への理解・教育の推進

図 10 医療 AI において解決されるべき諸課題

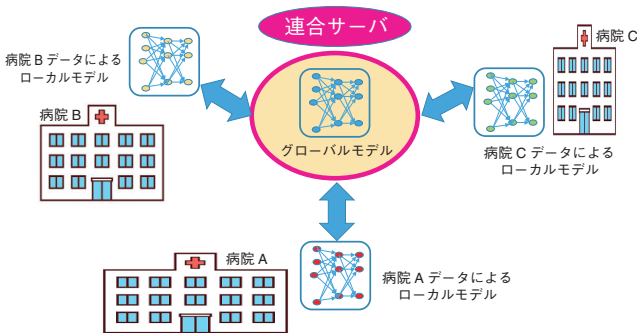


図 9 連合学習の概念図

も重要である。医用画像の世界では、一般の画像データ数と比べてきわめて少ないケースが多いが、それでもまず 1000 症例クラスでシステムを構築し、1 万症例規模でシステム性能を向上させ、次に 10 万症例規模で実用化を目指し、本格的な世界展開の商用化には 100 万症例規模で、という大雑把な印象がある。

医療画像のための公共のデータベースも増えてはいるが、多種多様にわたる各種診断領域をカバーはできていない。また、個人情報保護やセキュリティ対策の観点から、医用画像の収集は容易ではない<sup>2)-4)</sup>。さらには学習に用いる正解データを揃えること(アノテーション、ラベル付け)も大変な労力が必要となる。国内では、(国研)日本医療研究開発機構(AMED)主導の医療画像データ収集活動があり、また「次世代医療基盤法」<sup>2)3)</sup>が制定され、収集した医療情報をビッグデータとして分析・活用しやすくなる枠組みも進みだしているが、これらはまだ本格的に活用が進んでいるとはいえない。

そこで、基本技術のデータオーギュメンテーション(画像処理による画像データの水増し)や、転移学習(自然画像で学習済みモデルの医用画像への転用)あるいはファインチューニング(転移学習に加え、医用画像データでさらに学習を追加することによる精度向上)が多く利用されている<sup>1)-4)</sup>。最近の話題は、連合学習(federated learning)という方法で、各施設内で個別に分散学習させ(データの持ち出しはなし)、共通サーバー内でモデルのみを共有して、「連合モデル」に進化させるものである(図 9)。これ

らの諸技術は、十分なデータがまだ集積されていないような昨今の COVID-19 感染症の画像診断における AI-CAD でも、大いに活用されている。

## 6. 諸 課 題

このように期待が大きい AI-CAD ではあるが、解決すべき課題も多い(図 10)。

最近では、データの量のみではなく、異なる施設の画像データも含めないと、AI を取り扱った研究論文として認められないようになってきている学会もある。画像データの量と質、撮像装置の機種や撮像条件の相違、性別、人種などなど、ディープラーニングの性能に依存する多くのパラメータがあり、これらの制約を取り除かないと、実臨床現場では使えない。管理された研究環境よりも実臨床はもっと複雑で多様である。従来型 CAD でも同様であったが、実験室レベルでのいわゆる後ろ向き(retrospective)評価のみならず、多様な状況下での実臨床現場での AI-CAD の有効性評価が重要である。米国の北米放射線学会(RSNA)からは、実臨床での AI 評価に役立つ 9 つの重要な考慮リストを提示している<sup>11)</sup>。

AI-CAD の有効性を示す論文が多数増産されている昨今であるが、最近 Nature 誌に掲載された「AI の透明性と再現性」と題する論文では、そのような研究成果において、他の医療機関でも実証できるように、コンピュータコード、予測モデル、テスト設定、データなどを共有できるように要求している<sup>12)</sup>。

ディープラーニングの問題点の 1 つに、導き出された結果の理由を説明できない、というブラックボックス性がある。そこで、何とか理由を導き出しホワイトボックス化しようとの試みがなされている。例えば、Grad-CAM (gradient-weighted class activation mapping) と呼ばれる手法で、ディープラーニングが画像のどの部分に反応しているかを視覚的にヒートマップで表示する方法があり(図 11<sup>13)</sup>)、CAD が見当違いの箇所を異常の候補として指摘しているようなケースでは、一目瞭然に医師が判定できる。それでも、これだけではその根拠の十分な説明とはいえない。よって、「説明可能な AI」(explainable AI) の

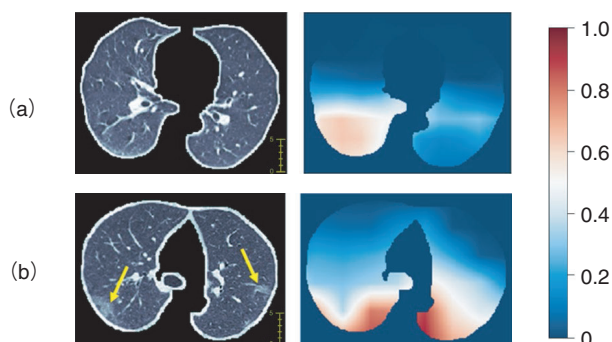


図 11 ディープラーニング (CNN) へ入力された CT 画像 (左) とその出力結果のヒートマップ表示 (右: 赤いほど, COVID-19 の感染確率が高い). (a) 発熱と感染歴のある 51 歳女性で, CNN モデルでは右下葉に異常 (右肺下端の白色部分) が認められたが, 2 名の放射線科医はこれを CT 陰性としたケース. (b) 感染歴があり発熱と湿性咳嗽 (しっせいがいそう: 痰のからんだ咳, ゴボゴボといった咳) を呈した 52 歳女性で, 両方の肺の辺縁部のすりガラス状の不透明度 (矢印) があり, 医師も CNN も正しく検出した例. 転載許諾を得て改変・引用<sup>13)</sup>. 図は J-STAGE ではカラー (冊子ではモノクロ) になっています.

開発が強く求められている。

また, 学習データ不足は医療分野では常に付きものであり, 深層学習の次なるイノベーションとして期待される正解データのないデータから特徴を抽出する「教師なし学習」や, あるいは学習データから正解を自ら作り出す「自己教師あり学習」などの斬新な技術開発が待たれる。

### 7. おわりに: 最終診断は医師

ディープラーニング技術の CAD への導入により, このように AI-CAD として, 従来の技術的な限界を突破して, 新しい画像支援診断の時代に突入している. AI-CAD により, 従来の CAD では達成できなかったような診断精度の向上がなされ, 多種疾病, 多種モダリティにも対応でき, 常時 ON でシームレスに利用できるようになりつつある. また, アプリ形式で必要な AI-CAD をダウンロードして (あるいはクラウド上で) 利用する枠組みもできつつある. 導入後もより賢く進化する AI-CAD が出現するのも時間の問題であろう。

AI-CAD の高度化・多様化の時代には, 「AI を理解して正しく使わないドクターは, AI を賢く利用するドクターに駆逐される」といわれるような昨今であり, 画像診断領域での AI の正しい理解と積極的な活用が望まれる。

最後に, 現状は「最終判断は医師によるものでなければならぬ」。2018 年 12 月, 厚生労働省からの通達「人工知能 (AI) を用いた診断, 治療等の支援を行うプログラムの利用と医師法第 17 条の規程との関係について」では, 「……診断, 治療等を行う主体は医師であり, 医師はその最終的な判断の責任を負うこと……」と注意を喚起してい

る. 10 年後にはもしかすると状況は変わっているかもしれないが, 現在はこれが鉄則である。

本解説では, 画像診断支援 (CAD) 領域における AI 導入の歴史から現状と課題について, 主に開発技術よりはむしろ臨床実利面から概説したが, 例として示した画像診断に限らず治療や手術支援など AI 応用領域は幅広い. より詳細な内容に興味ある読者は, 文献 2)-4), 14) を参照されたい. 本研究の一部は, 科研費・基盤研究 (B) (19H03599) による。

### 参考文献

- 1) 藤田広志: 第 3 次 AI ブーム・ディープラーニング時代における医用画像診断の最近の話題, 精密工学会画像応用技術専門委員会研究会報告, **35**, 4 (2020) 1.
- 2) 藤田広志 (編著): 医療 AI とディープラーニングシリーズ 2020-2021 年版 はじめての医用画像ディープラーニング基礎・応用・事例一, オーム社, 東京 (2020).
- 3) 藤田広志 (監): 学ぶ! 極める! 医療 AI—ディープラーニングの基礎から研究最前線まで一, インナービジョン, 東京 (2020).
- 4) 井川房夫, 藤田広志 (共編): これだけでわかる! 医療 AI, 中外医学社, 東京 (2021).
- 5) <https://labcoat.jp/world-ranking-for-ct-mri-pet/> accessed 2021.6.15.
- 6) J.D. Keen, J.M. Keen and J.E. Keen: Utilization of computer-aided detection for digital screening mammography in the United States, 2008 to 2016, *J. Am. Coll. Radiol.*, **15**, 1 Pt A (2018) 44.
- 7) J.J. Fenton et al.: Influence of computer-aided detection on performance of screening mammography, *N. Engl. J. Med.*, **356**, 14 (2007) 1399.
- 8) <https://www.youtube.com/watch?v=2HMpRXstSvQ>
- 9) A. Esteva et al.: Dermatologist-level classification of skin cancer with deep neural networks, *Nature*, **542**, (2017) 115.
- 10) H.A. Haenssle et al.: Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists, *Ann. Oncol.*, **29**, 8 (2018) 1836.
- 11) D.A. Bluemke et al.: Assessing radiology research on artificial intelligence: A brief guide for authors, reviewers, and readers—From the *radiology* editorial board, *Radiology*, **294**, 3 (2020) 487.
- 12) B. Haibe-Kains et al.: Transparency and reproducibility in artificial intelligence, *Nature*, **586** (2020) E14.
- 13) X. Mei et al.: Artificial intelligence-enabled rapid diagnosis of patients with COVID-19, *Nat Med.*, **26**, 8 (2020) 1224.
- 14) 藤田広志 (監), 有村秀孝, 諸岡健一 (編): 医療 AI とディープラーニングシリーズ 放射線治療 AI と外科治療 AI, オーム社, 東京 (2020).



藤田広志

岐阜大学大学院工学研究科修了. 工学博士 (名古屋大学). 岐阜高専, シカゴ大学を経て, 岐阜大学工学部/医学部教授. 現在, 岐阜大学特任教授/名誉教授. 専門: 医用画像の計算機支援画像診断学. 電子情報通信学会フェロー. 医用画像情報学会名誉会長. 医療 AI 関係などの著書多数.